SciencesPo
LABORATOIRE INTERDISCIPLINAIRE
D'ÉVALUATION DES POLITIQUES PUBLIQUES

LIEPP Working Paper

April 2019, nº87

Macroeconomic Imbalance Procedure, Economic Reforms and Policy Progress in the European Union

Jean-Charles BRICONGNE
European Commission, Banque de France, University of Tours
affiliate at Laboratory of Economics of Orléans and LIEPP
jean-charles.bricongne@banque-france.fr

Nuria MATA GARCIA
European Commission, Macroeconomic imbalances and adjustment Unit

Alessandro TURRINI
European Commission, Macroeconomic imbalances and adjustment Unit

www.sciencespo.fr/liepp

How to cite this publication:
BRICONGNE J.-C., MATA GARCIA N., TURRINI A., **Macroeconomic Imbalance Procedure, economic reforms and policy progress in the European Union**, *Sciences Po LIEPP Working Paper* n°87, 2019-04-15

Macroeconomic Imbalance Procedure, economic reforms and policy progress in the European Union [1]

Jean-Charles Bricongne[2,3,4] , Nuria Mata Garcia[5] and Alessandro Turrini[6]

Abstract

Every year, in the macroeconomic imbalances procedure (MIP), the European Commission examines the economic situation of member States, decides whether to launch an in-depth-review (IDR) and classifies countries into several categories, ranking from "no imbalances" to "excessive imbalances". The European Commission then releases some "specific country recommendations" (CSRs), detailing the economic measures to take to address the challenges and the imbalances.

This procedure has few equivalents in the world and, in that context, the question of the extent to which the pressure stemming from the MIP procedure can incite member States to implement reforms can be raised.

It is found that the pressure induced by the MIP classification is associated with more progress, whatever the regression used and whatever the controls.

Besides, if control variables' coefficients are not always significant depending on the regressions, their signs are as expected: difficult or politically/socially sensitive recommendations are associated with less progress, especially in the area of structural reforms and public finances. As regards political factors, progress is relatively less important when the mandates are getting close to their ends than for newly elected governments.

Keywords: structural reforms; macroeconomic imbalances procedure; macroeconomic surveillance; policy progress; international economic coordination.

JEL codes: E61, F42.

[1] The views expressed in this article are the ones of the authors and not the ones of their institutions of affiliation.

[2] European Commission, Office: CHAR 13-146 / B-1049 Brussels (Belgium), Rue de la Loi/Wetstraat 170. E-mail: jean-charles.bricongne@banque-france.fr

[3] Deputy-Director, Directorate of International Economics and Cooperation, Banque de France, rue Croix des Petits Champs, 75001 Paris, France

[4] Tours University, Laboratory of Economics of Orléans (LEO) and LIEPP (Laboratoire Interdisciplinaire d'Evaluation des Politiques Publiques, Sciences Po Paris).

[5] European Commission / B-1049 Brussels (Belgium), Rue de la Loi/Wetstraat 170.

[6] European Commission, Head of the Macroeconomic imbalances and adjustment Unit, Office: CHAR 13-173 / B-1049 Brussels (Belgium), Rue de la Loi/Wetstraat 170. E-mail: Alessandro.TURRINI@ec.europa.eu

1. Introduction and related literature

The experience of the Global Crisis required policymakers to broaden EU macroeconomic surveillance beyond fiscal aspects. A number of macroeconomic imbalances (relating, for example, to current account balances, competitiveness divergences, private indebtedness, housing market dynamics, etc.) revealed themselves as key factors in triggering balance-of-payment crises and the recourse to financial assistance. The Macroeconomic Imbalance Procedure (MIP) was introduced in 2011 with the aim of preventing the accumulation of harmful macroeconomic imbalances and foster their correction, once in place, by means of a system of surveillance comprising of recommendations and possible sanctions.[7]

MIP surveillance follows a regular cycle. Cross-country analysis carried out in an Alert Mechanism Report (AMR) provides the basis for selecting a number of countries that also would be analysed in in-depth-reviews (IDR) by the Commission, with a view to assess the existence of imbalances that are harmful for macroeconomic stability, and evaluate their severity.[8]

The in-depth reviews may result in the identification of no imbalances, imbalances, or excessive imbalances. Countries identified as having imbalances or excessive imbalances receive Country-Specific Recommendations (CSRs) by the Commission and the EU Council in the context of the 'European semester'. For the countries identified with excessive imbalances, the Excessive Imbalance Procedure (EIP) may also be activated, which comprises of the delivery of a corrective action plan with a set of policy measures to be carried out within a pre-determined time frame. The repeated delivery of an insufficient corrective action plan or repeated lack of compliance with the policy measures detailed in the plan may imply sanctions for the countries that belong to the Eurozone.

The two arms – preventive and corrective – similarly to those in the Stability and Growth Pact (SGP) have different objectives: the preventive arm helps Member States to adopt good policies that will lead to balanced medium-term growth, jobs and financial stability. The corrective arm aims to identify and correct policy failures or address major macroeconomic risks that are harmful for the economic developments in the Member State concerned and may generate negative spillovers to other Member States.

Since its inception, MIP surveillance has been activated mainly to foster an orderly correction of imbalances accumulated before the financial crisis. In some cases, MIP surveillance was replaced by financial assistance programmes (Spain and Cyprus). In other situations, MIP surveillance was activated soon after the conclusion of programmes (Spain, Ireland, Romania, Portugal, and Cyprus).

So far, the EIP has not been launched. The identification of excessive imbalances was instead followed by relatively prescriptive recommendations with deadlines, enhanced policy commitments by the Member States concerned (contained in their National Reforms Programmes), and a process of "specific monitoring" of the implementation of MIP-related policy commitments. The

[7] The legal basis of the MIP are two regulations: one outlining the steps in the procedure (regulation (EU) 1176/2011), and one providing a scheme of sanctions in case of reiterated lack of compliance by Eurozone Member States (regulation (EU) 1174/2011).

[8] In Regulation (EU) 1176/2011 imbalances are given a relatively broad definition, being "any trend giving rise to macroeconomic developments which are adversely affecting, or have the potential adversely to affect, the proper functioning of the economy of a Member State or of the economic and monetary union, or of the Union as a whole".

categorisation of MIP imbalances by the Commission has also been more articulated than envisaged by the MIP regulation, including considering the necessity of policy action and monitoring.

Was the MIP effective in triggering policy action despite that fact that no use was made of the enhanced surveillance framework in the EIP Procedure, and the procedure remained far from the step of sanctions?

Similar questions have been raised in recent studies. Darvas and Leandro (2015) analyse compliance to the Commission CSRs. They find a higher level of implementation for CSRs that are related to the Stability and Growth Pact (SGP), followed by MIP-related CSRs, and then the remaining CSRs. A straightforward interpretation would be that the former two sets of CSRs are backed by a relatively higher degree of enforcement, as the associated procedures are the only ones foreseeing the possibility of sanctions. However, results may depend on other factors as well, including a discipline role played by markets or the political capital absorbed by complying to CSRs. The present analysis carries controls for these factors.

It is the first time an assessment of the influence of a common procedure for the coordination of macroeconomic policies between the member States of an integrated area (namely the European Union) has been made, using the procedures launched in the aftermath of the crisis, to foster the correction of domestic and international imbalances.

The Stability and Growth Pact has already been reviewed and criticized, with some ways for improvement proposed (Cf. for example Schuknecht et al. (2011) or Fischer et al. (2006)), but without any quantitative assessment on progress made by itself.

The main quantitative evaluations that have been made are connected to the potential macroeconomic impact rather than on the assessment of progress of effective actions. Creel et al. (2012) is an example for such a simulation. This paper discusses the different reforms and subsequent fiscal rules which have emerged since 2011. It assesses the impact of fiscal rules on the output gap and inflation rate of three representative countries of the Eurozone, using a counterfactual. Castro (2011) also analyses whether Maastricht and Stability and Growth Pact fiscal rules have affected growth in the European Union negatively. A growth equation is specified for a group of 15 European Union countries (and 8 OECD countries) over the period 1970–2005 to analyse this issue. It is shown that the institutional changes that occurred in the European Union after 1992 were not harmful to growth and that growth was even slightly higher in the period in which the fulfilment of the 3% criteria for the deficit started to be officially assessed, i.e. after 1997.

When looking at the role of external inducements like IMF programs, Alesina et al. (2006) find that they have at best a weak effect, and underline that problems of reverse causality are possible, in the sense that foreign aid should go to countries in trouble, so a correlation of bad policies and delayed stabilizations with foreign aid can have different causal interpretations. Another pessimistic view about the effects of foreign aid is in Easterly (2006).

In the case of the influence of the MIP procedure, these problems of reverse causality should be limited in the sense that the assessment of progress made induces no granting of foreign aid which may incite to deter or defer reforms, only incentives to reform further.

Moreover, the impact of the classification in the MIP procedure is introduced with one-year lag, compared to progress assessment, thus reverse causality is less likely.

In addition, reverse causality is all the less likely to happen since more progress induces on the contrary an easing of the classification at the end of rolling year, as shown in the Annex 2 of European Commission (2016), controlling besides for a "synthetic indicator of economic conditions", capturing the relevant variables summarizing the imbalances identified in IDRs for related countries.

This effect goes in the opposite direction of the one tested, and thus pleads that coefficients' variables, if there is to be reverse causality should be under-stated.

The problem would rather be the one of omitted variable, which would be correlated with the MIP classification. To try to face this problem, several key control variables (markets pressure, GDP growth, electoral cycle) are introduced, but possible biases may still be observed. Hence our results should rather be interpreted as correlations at this stage; they do not capture causal effects.

The second part analyses the data used and give some stylized facts about CSRs and MIP categories. The third part will perform econometric relations, taking into account the impact of the MIP category and controlling for other factors, including potential reform fatigue; complementary regressions address the issue of stability of relations over time, potential endogeneity and selection bias. The last part concludes.

2. Data used and stylized facts

2.1. Description of the variables used

As regards the dependent variable, which sets policy progress as assessed by the European Commission, each data point corresponds to recommendations at a disaggregated policy field, defined at a more detailed level than that found in Council CSRs as published (namely "sub-CSRs").

Progress with respect to recommendations is thus defined at the following level of disaggregation (for more detail, see Annex): Public finances; Financial sector; Labour market, social inclusion and education; Structural policies; and Public administration and business environment.

The indicator of policy progress distinguishes five degrees of progress, ranging from 0 to 100: 0=no progress; 25=limited progress; 50=some progress; 75=substantial progress; and 100=full achievement.

The main explanatory variable that is tested is the classification of each country in a MIP category, the year before progress is assessed, which signals the magnitude of the pressure induced by the European Commission. There are three categories for MIP imbalances (no imbalances=0; imbalance=1; excessive imbalances=2).

Besides this main explanatory variable, several key control variables are envisaged (see their definition in Annex), which are likely to influence the extent to which countries may comply with policy recommendations.

A dummy ("hard-to-comply with recommendation") takes value one for policy fields where the political cost of reforms is higher either because the group of potential losers is wide or because it

is well organised (e.g. Olson (1965)). This if important to control for since progress to make may be more difficult if the related sub-CSRs fall in this category. This variable is partly judgemental, but a robustness regression is performed without this indicator (Cf. table 3, columns [5] and [6] in the sub-part 3.1).

GDP growth enables to control for the current economic situation, specific to the country (since there is besides a year fixed effect, which controls for the global economic situation). This variable may potentially have two different impacts. First, if the economic situation is favourable, it may deter the government to reform, waiting for the positive effects of economic growth. Conversely, it may help to smooth the potentially negative short-term impact of reforms in terms of social and economic costs. Thus, the overall effect is undetermined, but this variable should be controlled for in any case.

The sovereign spread with the benchmark German rate indicates the potential pressure imposed by financial markets. It is lagged to avoid endogeneity problems.

To control for the political cycle, dummies indicating legislative elections at most one year before, or by next year, are introduced, using the World Bank political database. Indeed, Alesina et al. (2006), analysing why countries delay stabilizations of large and increasing budget deficits and inflation, on a vast sample of countries, find that stabilizations are more likely to occur when time of crisis occur, at the beginning of term of office of a new government, in countries with "strong" governments (i.e. presidential systems and unified governments with a large majority of the party in office), and when the executive faces less constraints. The nature of political regimes (presidential or not) having a limited variance on the short period considered in the article, the most important factor to control for is thus the fact of being or not at the beginning of term of office, which is done by considering recent elections and elections to come.

The policy area to which the sub-CSR belongs are split between the five categories detailed above (public finances; financial sector; labour market, social inclusion and education; structural policies; Public administration and business environment). This is introduced as a control variable since average progress may not be the same depending on the policy area. For example, structural policies are expected to be difficult to implement (e.g. Rajan (2004)), in spite of their potential impact on productivity (e.g. Nicoletti and Scarpetta (2003) or Bourlès et al. (2013) for a panel of OECD countries).

Besides these standard factors, in the first two columns, to capture potential reform fatigue, past progress and number of sub-CSRs, cumulated until the year before, have been taken into account. More precisely, for each year, countries are classified by quartiles, depending on the sum of progress over all their sub-CSRs and the number of sub-CSRs, higher quartiles corresponding respectively to more progress and more sub-CSRs. This calculation enables to neutralize the streamlining of the number of CSRs and sub-CSRs which appears in the second part of the period under review, and have homogenous calculations.

2.2. Stylized facts

In the whole, 23 countries are covered, namely all the member States of the European Union (except Croatia, Cyprus, Greece, Ireland and Portugal, which were under program and/or not covered by the MIP procedure), over the years 2013 to 2017.

As can be seen in table 1, the number of sub-CSRs is more important for countries with excessive imbalances, compared with countries with imbalances or no imbalances, the latter two categories displaying limited difference.

Besides, from 2015, the number of sub-CSRs diminishes compared to previous years, due to a "streamlining" of recommendations, which induces some heterogeneity over time that must be taken into account when considering reform fatigue (see the definition of cumulated number of sub-CSRs in Annex).

Table 1: **Number of sub-CSRs per MIP classification**

Year	MIP status	Median sum of sub-CSRs by country	Mean sum of sub-CSRs by country
2013	No imbalance	21,0	20,2
2014	No imbalance	17,0	16,4
2015	No imbalance	9,0	8,3
2016	No imbalance	8,5	8,4
2017	No imbalance	8,0	7,1
2013	Imbalance	19,0	18,4
2014	Imbalance	17,0	20,9
2015	Imbalance	10,0	9,3
2016	Imbalance	11,0	8,9
2017	Imbalance	8,5	8,2
2013	Excessive imbalance	31,0	31,0
2014	Excessive imbalance	26,5	26,5
2015	Excessive imbalance	17,5	16,3
2016	Excessive imbalance	16,0	15,0
2017	Excessive imbalance	13,5	13,3

Sources: European Commission, authors' calculations.

When analysing the average progress by MIP classification and year (see table 2), countries under the "excessive imbalance" classification display more progress compared to the other two categories, at least until 2015. Then, the difference falls in 2016 and to a lesser extent in 2017, which suggests, *prima facie*, some possible reform fatigue in the implementation of recommendations and reforms.

Table 2: **Average progress per MIP classification**

Year	MIP status	Average progress
2013	No imbalance	37,1
2014	No imbalance	37,2
2015	No imbalance	36,3
2016	No imbalance	39,2
2017	No imbalance	35,8
2013	Imbalance	39,5
2014	Imbalance	35,7
2015	Imbalance	38,9
2016	Imbalance	35,3
2017	Imbalance	36,0
2013	Excessive imbalance	45,2
2014	Excessive imbalance	49,5
2015	Excessive imbalance	41,5
2016	Excessive imbalance	39,0
2017	Excessive imbalance	39,6

Sources: European Commission, authors' calculations.

3. Econometric results: the MIP procedure has a positive impact, structural and public finances reforms are more difficult to implement and reform fatigue has nuanced effects

3.1. Baseline specifications

Assessing to what extent MIP surveillance strengthens policy progress requires controlling for other factors that play a role in driving reform outcomes. To this purpose, a regression specification has been tested where an indicator of progress with respect to previous EU recommendations is put in relation with the categories for MIP imbalances and additional control variables that are likely to influence the extent to which countries are likely to comply with policy recommendations.

Taking into account all the explanatory variables detailed in the sub-part 2.1, the baseline specification is as follows:

Log (1+sub-CSR compliance score) = α_1.constant + α_2.MIP imbalance category + α_3.hard-to-comply with recommendation dummy + α_4.GDP growth + α_5.interest rates spreads in previous year + α_6.Elections in previous year + α_7.Elections by following year + α_8.cumulated progress + α_9.cumulated number of sub-CSRs + v_t + + v_{pa} + ε_{ct}

With v_t = year fixed effects, v_{pa} dummies for policy areas and ε_{ct} the residual of the equation.

As regards the dependent variables, the scores are transformed as log(1+score) to smooth the effect of the retained scaling, which ranks from 0 to 100, and may otherwise be somewhat arbitrary.

Since the dependent variable is detailed at the country / year / sub-CSR level and the explanatory variables at the country / year level, clustering at the country / year level has been performed.

Moreover, the explanatory variable being not continuous, ordered probit are performed in columns [1], [3] and [5], and ordinary least squares (OLS) in columns [2], [4] and [6] of table 3 as robustness checks.

Results displayed in table 3 show that in all cases, over the whole period, the MIP imbalance category has a significant and positive impact on progress, at the 5% level. with approximately the same magnitude. This suggests that MIP surveillance favours indeed policy progress and seems to be robust.

Recommendations which are hard-to-comply with, *a priori*, have a negative and significant impact at the 1% level (columns [1] to [4]). This is as expected, since these recommendations are more difficult to implement, being politically costly. The omission of this variable in columns [5] and [6] does not alter main results.

GDP growth and interest rate spreads are not significant. Higher growth should make reforms easier to implement, despite potential complacency. The fact that the opposite effect may also play (namely a moral hazard effect for governments which are less prone to reform when relying on GDP growth) may explain the overall low significance of this variable.

The elections to come are always significant and have a relatively more negative impact on progress than recent elections. This is consistent with the usual findings of electoral cycles (Cf. for example Alesina et al. (2006)).

The past cumulated number of sub-CSRs plays indeed a negative and significant role at the 1% level, which means that the more reforms have been required in the past, the less progress is implemented, which would be consistent with the idea of reform fatigue. This being said, if the country has cumulated high progress in the past, it is more likely to go on doing so, the variable "cumulated progress until previous year" being significant at the 1% level and positive. Thus, this dual finding nuances the idea of reform fatigue.

In terms of policy areas, public finances and structural policies are the ones associated with least progress, which is in line with the priors associated to difficulties of these policies. It is in particular the case for columns [5] and [6], where the “hard-to-comply with recommendation” variable, which captures in particular components of structural reforms’ policy area (see definition in Annex) is not included.

Table 3: **MIP surveillance and CSR progress, data at detailed recommendation level, all EU members (except countries under programme)**

Dependent variable: CSR compliance score	[1]	[2]	[3]	[4]	[5]	[6]
Explanatory variables						
Estimation method:	Ordered probit	OLS	Ordered probit	OLS	Ordered probit	OLS
MIP imbalance category	0.175**	0.151**	0.153**	0.135**	0.149**	0.132**
	(0.074)	(0.068)	(0.068)	(0.061)	(0.067)	(0.061)
Hard-to-comply with recommendation, dummy	-0.241***	-0.284***	-0.205***	-0.250***		
	(0.069)	(0.089)	(0.068)	(0.088)		
GDP growth	-0.013	-0.035	0.017	-0.007	0.016	-0.009
	(0.030)	(0.032)	(0.030)	(0.032)	(0.030)	(0.032)
Interest rate spread in previous year	0.020	0.012	-0.012	-0.012	-0.009	-0.008
	(0.028)	(0.031)	(0.029)	(0.033)	(0.029)	(0.033)
Cumulated progress until previous year	0.090***	0.080***				
	(0.021)	(0.021)				
Cumulated number of sub-CSRs until previous year	-0.067***	-0.055***				
	(0.019)	(0.018)				
Elections in previous year	-0.077	-0.096	-0.165*	-0.168*	-0.157*	-0.159*
	(0.088)	(0.087)	(0.095)	(0.096)	(0.094)	(0.096)
Elections by following year	-0.199**	-0.217**	-0.224**	-0.246**	-0.219**	-0.240**
	(0.100)	(0.102)	(0.108)	(0.111)	(0.108)	(0.112)
Policy area “financial sector”	0.087	-0.149	0.131	-0.103	0.184	-0.038
	(0.192)	(0.202)	(0.184)	(0.199)	(0.179)	(0.197)
Policy area “labour market, social inclusion & education”	-0.144	-0.054	-0.161*	-0.058	-0.172*	-0.073
	(0.091)	(0.110)	(0.090)	(0.108)	(0.090)	(0.109)
Policy area “public administration & business environment”	0.099	0.179	0.048	0.137	-0.046	0.024
	(0.116)	(0.138)	(0.110)	(0.132)	(0.111)	(0.123)

Policy area "structural policies"	-0.139	-0.043	-0.158*	-0.064	-0.251***	-0.178*
	(0.091)	(0.121)	(0.086)	(0.112)	(0.087)	(0.104)
Policy area "public finances"	-0.164*	-0.304***	-0.162*	-0.299***	-0.177*	-0.319***
	(0.093)	(0.114)	(0.092)	(0.112)	(0.091)	(0.112)
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Clustering at country-year level	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1463	1463	1495	1495	1495	1495
Pseudo R-squared	0.03		0.02		0.01	
R-squared		0.05		0.03		0.03
RMSE		1.23		1.23		1.24

Note: *, ** and ***: coefficients significant at the 10%, 5% and 1% level respectively. Student-t are reported in parenthesis. Standard errors are robust with-respect to clustering at the country-year level.
Dependent variable: CSR compliance score defined at disaggregated policy field. 0=no progress; 25=limited progress; 50=some progress; 75=substantial progress; 100=full achievement. The scores are transformed as log(1+score) to smooth the effect of the retained scaling.
MIP categories: no imbalances=0; imbalances=1; excessive imbalances=2. This detail has been retained to fit with the MIP categories in the end of the period. The MIP categorisation refers to the preceding year. The correspondence between the classifications in MIP categories used in 2013 and 2014 is the one displayed in table 4.3 of European Commission (2016), with the two categories qualified as being excessive imbalances in 2013 corresponding to excessive imbalances category of 2014.
The "hard-to-comply with recommendations" variable is a dummy variable taking value 1 for sub-CSRs belonging to the following fields: Long-term sustainability of public finances, including pensions; Employment protection legislation & framework for labour contracts; Wages & wage setting; Health & long-term care; Competition in services; Telecom, postal services & local public services; Energy, resources & climate change; Transport; Public administration; State-owned enterprises; Civil justice; Shadow economy & corruption.
Interest rate spreads are expressed as the difference of the 10-year government bond yield with the one of Germany.
"Elections in previous year" and "Elections in following year" are dummies equal to one if legislative elections have taken place at most one year before the related year or if these elections are to take place in at most one year, respectively.
Cumulated progress and number of sub-CSRs until previous year are based on classifications by quartiles, depending on the sum of progress over all sub-CSRs and the number of sub-CSRs, higher quartiles corresponding respectively to more progress and more sub-CSRs.
Source: European Commission, Eurostat, World Bank political database, national sources, own calculations.

3.2. Robustness checks: instrumental regressions confirm the impact of the MIP procedure; results seem to alter in the end, but not for countries remaining in the procedure

The VOX column by Bricongne and Turrini (2017), elaborating on the results from European Commission (2016) finds some impact of the MIP procedure on progress over the first years, without implementing the excessive imbalance procedure or any sanctions. But the relation shown in Table 3 may still alter over time. Indeed, some recent documents such as Efstathiou and Wolff (2018) seem to challenge the impact of the MIP procedure: "Overall implementation of recommendations by EU countries has worsened in the last few years, in particular when it comes to recommendations addressed to countries with excessive macroeconomic imbalances", as could also be noticed in the statistics in the sub-part 2.1.

Still, the impact of the MIP procedure should be established with specifications that enable to control for alternative indicators to avoid the bias of omitted variables (as already done in Table 3), controlling also for potential reverse causality and selection biases.

Complementary regressions have been performed to address these issues in table 4.

To check for reform fatigue, besides what is done in columns [1] and [2] in table 3, the period is split into two (see columns [1] and [2] of table 4).

Since progress may influence MIP categorisation, as shown in European Commission (2016), in Annex 2, there may be reverse causality, even if the influence should be in the opposite direction (more progress will improve the categorisation and thus diminish the value of MIP category, being given the retained convention in calculations). To neutralize this source of potential endogeneity, instrumentation is performed in columns [3] and [4] of table 4, over the whole period and the last two years respectively. Instruments correspond to the synthetic indicators of economic conditions in stocks and flows, used in Annex 2 of European Commission (2016) (see also definition in Annex), taking into account imbalances identified as affecting each country in in-depth reviews.

Besides, since countries that exit the procedure are, by definition, those which have implemented more progress, there may be a selection bias. To control for this effect, a sub-sample of countries which have been classified with imbalances or excessive imbalances over the whole period has been retained in columns [5] and [6], corresponding to France, Italy, Bulgaria, Netherlands, Spain and Sweden.

Results in table 4 show that the MIP procedure impact seems to vanish in the end of the period (column [2]). Instrumentation over the whole period confirms and even amplifies the magnitude of the results for the impact of the MIP procedure, the Fisher coefficient checking the rule of thumb of 10 (column [3] compared to column [1])[9] but significance disappears again on the end of the period (column [4]).

Still, when concentrating on the sub-sample of countries which are always classified with imbalances, this impact remains, whatever the period (columns [5] and [6]).
As regards the difficulty of implementation of structural policies, this policy area is significant at the 5% level and negative for this sub-sample, the "public finances" policy area and the "hard-to-comply with recommendation" variable turning non-significant. The latter remains significant at the 1% or 5% level and negative whatever the sub-period (columns [1] and [2]) and when instrumentation is implemented (columns [3] and [4]).

[9] The instrument used is the synthetic indicator of economic conditions (see definition in Annex), based among other variables on general government debt. Using this single indicator alone as instrument confirms the results. The MIP imbalance category coefficient is even higher and equal to 0.441 and still significant (at the 10% level though) with the Fisher coefficient equal to 9.9. Kriesi (2014) confirms the political impact of this variable.

Table 4: **MIP surveillance and CSR progress, complementary specifications using sub-periods, instrumentation and sub-samples of countries always with imbalances**

	Sub-period [2013:2015]	Sub-period [2016:2017]	MIP imbalance category instrumented: [2013:2017]	MIP imbalance category instrumented: [2016:2017]	Sub-sample of countries always with imbalances: [2013:2017]	Sub-sample of countries always with imbalances: [2016:2017]
Dependent variable: CSR compliance score	[1]	[2]	[3]	[4]	[5]	[6]
Explanatory variables						
Estimation method:	Ordered probit	Ordered probit	IV Regress	IV Regress	Ordered probit	Ordered probit
MIP imbalance category	0.189**	0.088	0.244**	0.057	0.501***	0.706**
	(0.092)	(0.102)	(0.100)	(0.214)	(0.150)	(0.357)
Hard-to-comply with recommendation, dummy	-0.179**	-0.385***	-0.289***	-0.403***	-0.051	-0.371
	(0.081)	(0.143)	(0.088)	(0.144)	(0.119)	(0.416)
GDP growth	0.003	-0.107	-0.018	-0.123	0.169***	0.042
	(0.033)	(0.089)	(0.033)	(0.087)	(0.063)	(0.095)
Interest rate spread in previous year	0.023	0.030	0.005	0.013	-0.020	0.124
	(0.031)	(0.064)	(0.033)	(0.060)	(0.047)	(0.188)
Cumulated progress until previous year	0.092***	0.107***	0.078***	0.091***	0.044	0.087
	(0.028)	(0.028)	(0.021)	(0.024)	(0.040)	(0.109)
Cumulated number of sub-CSRs until previous year	-0.053**	-0.108***	-0.059***	-0.068**	0.019	-0.087
	(0.025)	(0.026)	(0.020)	(0.031)	(0.041)	(0.084)
Elections in previous year	-0.072	0.101	-0.107	-0.002	-0.479**	0.481
	(0.113)	(0.141)	(0.086)	(0.142)	(0.191)	(0.305)
Elections by following year	-0.200	0.155	-0.238**	0.008	-0.562***	0.865***
	(0.124)	(0.200)	(0.105)	(0.192)	(0.195)	(0.336)
Policy area “financial sector”	0.087	0.130	-0.188	-0.022	-0.010	0.092
	(0.223)	(0.304)	(0.201)	(0.300)	(0.345)	(0.439)
Policy area “labour market, social inclusion & education”	-0.135	-0.084	-0.044	0.097	-0.204	-0.324

	(0.108)	(0.193)	(0.109)	(0.226)	(0.168)	(0.346)
Policy area "public administration & business environment"	0.088	0.164	0.183	0.272	-0.096	-0.061
	(0.144)	(0.194)	(0.138)	(0.264)	(0.249)	(0.301)
Policy area "structural policies"	-0.127	-0.259	-0.029	-0.133	-0.332**	-0.712**
	(0.103)	(0.209)	(0.121)	(0.239)	(0.137)	(0.289)
Policy area "public finances"	-0.158	-0.195	-0.293***	-0.248	-0.024	0.149
	(0.103)	(0.203)	(0.113)	(0.226)	(0.215)	(0.340)
Year F.E.	Yes	Yes	Yes	Yes	Yes	Yes
Control for policy area	Yes	Yes	Yes	Yes	Yes	Yes
Clustering at country-year level	Yes	Yes	Yes	Yes	Yes	Yes
Number of observations	1063	381	1514	381	449	117
Pseudo R-squared	0.02	0.05			0.06	0.06
R-squared			0.05	0.09		
RMSE			1.21	1.08		
Fisher coefficient (first stage of instrumentation)			73.2	21.3		

Note: *, ** and ***: coefficients significant at the 10%, 5% and 1% level respectively. Student-t are reported in parenthesis. Standard errors are robust with-respect to clustering at the country-year level.
Dependent variable: CSR compliance score defined at disaggregated policy field. 0=no progress; 25=limited progress; 50=some progress; 75=substantial progress; 100=full achievement. The scores are transformed as log(1+score) to smooth the effect of the retained scaling.
MIP categories: no imbalances=0; imbalances=1; excessive imbalances=2. This detail has been retained to fit with the MIP categories in the end of the period. The MIP categorisation refers to the preceding year. The correspondence between the classifications in MIP categories used in 2013 and 2014 is the one displayed in table 4.3 of European Commission (2016), with the two categories qualified as being excessive imbalances in 2013 corresponding to excessive imbalances category of 2014.
The "hard-to-comply with recommendations" variable is a dummy variable taking value 1 for sub-CSRs belonging to the following fields: Long-term sustainability of public finances, including pensions; Employment protection legislation & framework for labour contracts; Wages & wage setting; Health & long-term care; Competition in services; Telecom, postal services & local public services; Energy, resources & climate change; Transport; Public administration; State-owned enterprises; Civil justice; Shadow economy & corruption.
Interest rate spreads are expressed as the difference of the 10-year government bond yield with the one of Germany.
"Elections in previous year" and "Elections in following year" are dummies equal to one if legislative elections have taken place at most one year before the related year or if these elections are to take place in at most one year, respectively.
Cumulated progress and number of sub-CSRs until previous year are based on classifications by quartiles, depending on the sum of progress over all sub-CSRs and the number of sub-CSRs, higher quartiles corresponding respectively to more progress and more sub-CSRs.
The instruments used are the synthetic indicators of economic conditions (see definition in Annex).
Source: European Commission, Eurostat, World Bank political database, national sources, own calculations.

4. Conclusion

The pressure induced by the MIP classification is associated with more progress, and this result is robust whatever the regression used and whatever the controls. In this paper, we have tried to address potential endogeneity bias by using instrumentation techniques. However, there may still be biases linked to omitted variables in spite of the various indicators used. Hence our coefficients should rather be interpreted as correlations at this stage, not as causal effects.

As regards reform fatigue, the effects are nuanced: less progress seems to be observed in the end of the period but countries that have cumulated most progress in the past, controlling for the number of reforms to implement, and those that remain in the procedure, do not register such a fall.

This pressure, besides influences from financial market and business or political cycles, seems thus to promote additional efforts by member States or, at least, to concentrate these efforts in the required direction.

If controls are not always significant depending on the regressions, their signs are as expected: difficult or politically/socially sensitive sub-CSRs are associated with less progress and structural reforms and public finances seem to be the most difficult policy areas. As regards political factors, progress is relatively less important when the mandates are getting close to their ends than for newly elected governments.

Conversely, more progress induces an easing of the classification at the end of the period, as shown in the Annex 2 of European Commission (2016). There is thus a mutual influence between the two variables: a worsening of the MIP classification fosters more progress which, in turn, enables to ease subsequently the severity of the MIP classification.

Besides, the assessment of the MIP procedure is not complete yet, in the absence of the launching of the excessive imbalances procedure and of the corrective arm. If used, the influence and magnitude of this part of the procedure, compared to the preventive arm, would be interesting to estimate.

References

Alesina A., Ardagna S. and Trebbi F. (2006), "Who adjusts and when? On the political economy of reforms", NBER Working Paper", n°12049, February.

Bourlès R., Cette G., Lopez J., Mairesse J. and Nicoletti G. (2013), Do Product Market Regulations In Upstream Sectors Curb Productivity Growth? Panel Data Evidence For OECD Countries, The Review of Economics and Statistics, MIT Press, vol. 95(5), pp. 1750-1768, December.

Bricongne J.-C. and Turrini A. (2017), The EU Macroeconomic Imbalance Procedure: Some impact and no sanctions, VOX column, 22th June.

Castro V. (2011), "The impact of the European Union fiscal rules on economic growth", Journal of Macroeconomics, Volume 33, Issue 2, pp. 313-326, June.

Creel, J., Hubert, P. and Saraceno, F. (2012), "The European Fiscal Compact: A Counterfactual Assessment," Journal of Economic Integration, Center for Economic Integration, Sejong University, vol. 27, pp. 537-563.

Darvas, Z, and A Leandro (2015), "The limitations of policy coordination in the euro area under the European Semester" Bruegel Policy Contribution, Issue 2015/19.

Easterly, W. (2006), "The White Man's Burden", Penguin, New York.

Efstathiou K. and Wolff G. (2018), Is the European Semester effective and useful? Bruegel, Policy Contribution Issue, n°9, June.

European Commission (2016), The Macroeconomic Imbalance Procedure - Rationale, Process, Application: A Compendium, Institutional Paper 039, November.

Fischer, J., Jonung, L. and Larch, M. (2006), "101 proposals to reform the Stability and Growth Pact. Why so many? A Survey", European Economy – Economic Papers, No 267, European Commission, December.

Hauptmeier, S., Sanchez-Fuentes, A.J. and Schuknecht, L. (2011), "Towards expenditure rules and fiscal sanity in the euro area", Journal of Policy Modelling, Vol. 33, Issue 4, pp. 597-617, July-August.

Kriesi H. (2014), The Political Consequences of the Economic Crisis in Europe: Electoral Punishment and Popular Protest, Chapter 10 in Mass Politics in Tough Times: Opinions, Votes and Protest in the Great Recession, edited by Nancy Bermeo and Larry M. Bartels. Oxford: Oxford University Press, pp. 297-333.

Nicoletti G. and Scarpetta S. (2003), Regulation, productivity and growth: OECD evidence, Economic Policy, CEPR; CES; MSH, vol. 18(36), pp. 9-72, April.

Olson, M (1965), The Logic of Collective Action: Public Goods and the Theory of Groups, Cambridge MA: Harvard University Press.

Rajan R. (2004), Why are structural reforms so difficult, Finance and Development, IMF.

Schuknecht L., Moutot P., Rother P. and Stark J. (2011), The Stability and Growth Pact: Crisis and Reform, ECB Occasional Paper nº129, March.

ANNEX:
Definition of variables

Cumulated number of sub-CSRs until previous year: based on classifications by quartiles, depending on the number of sub-CSRs, higher quartiles corresponding to more sub-CSRs. Yearly quartile values are then cumulated from the start until previous year.

Cumulated progress until previous year: based on classifications by quartiles, depending on the sum of progress over all sub-CSRs, higher quartiles corresponding to more progress. Yearly quartile values are then cumulated from the start until previous year.

Elections in previous year: dummy equal to one if legislative elections have taken place at most one year before the related year (source: World Bank political database).

Elections in following year: dummy equal to one if legislative elections are to take place in at most one year from the related year (source: World Bank political database).

GDP growth: yearly growth rate of GDP at market price, chained-linked volume (source: Eurostat).

Hard-to-comply with recommendation dummy: The classification of political or social sensitivity is made by setting a dummy equal to one for the reforms which are most likely to generate social or political opposition (with likely strikes, strong political oppositions or social unrest). It corresponds to the following fields: Long-term sustainability of public finances, including pensions; Employment protection legislation & framework for labour contracts; Wages & wage setting; Health & long-term care; Competition in services; Telecom, postal services & local public services; Energy, resources & climate change; Transport; Public administration; State-owned enterprises; Civil justice; Shadow economy & corruption.

Interest rate spreads: difference of the 10-year government bond yield with the one of Germany. They are introduced with a one-year lag to minimize endogeneity issues.

Policy areas: the five areas correspond to the following fields:

- Public finances: Fiscal policy and fiscal governance; Long-term sustainability of public finances, including pensions; Reduce the tax burden on labour; Broaden tax bases; Reduce the debt bias; Fight against tax evasion, improve tax administration & tackle tax avoidance.
- Financial sector: Financial services; Housing market; Access to finance; Private indebtedness.
- Labour market, social inclusion and education: Employment protection legislation & framework for labour contracts; Unemployment benefits; Active labour market policies; Incentives to work, job creation, labour market participation; Wages & wage setting; childcare; Health & long-term care; Poverty reduction & social inclusion; Education; Skills & life-long learning.
- Structural policies: Research & innovation; Competition & regulatory framework; Competition in services; Telecom, postal services & local public services; Energy, resources & climate change; Transport.

- Public administration and business environment: Business environment; Insolvency framework; Public administration; State-owned enterprises; Civil justice; Shadow economy & corruption.

Synthetic indicator of economic conditions: this variable can be calculated using "stocks" or "flows". The one using "stocks" notions is a sum of the normalized values of the following variables: general government gross debt, financial sector leverage, NIIP (net international investment position), corporate debt and households' debt. The one using "flows" notions is a sum of the normalized values of the following variables: export market shares, unit labour costs, house prices, current account, financial sector liabilities' growth and REER (real effective exchange rate).

Le LIEPP (Laboratoire interdisciplinaire d'évaluation des politiques publiques) est un laboratoire d'excellence (Labex).
Ce projet est distingué par le jury scientifique international désigné par l'Agence nationale de la recherche (ANR).
Il est financé dans le cadre des investissements d'avenir.

(ANR-11-LABX-0091, ANR-11-IDEX-0005-02)

www.sciencespo.fr/liepp

A propos de la publication

Procédure de soumission :
Rédigé par un ou plusieurs chercheurs sur un projet en cours, le *Working paper* vise à susciter la discussion scientifique et à faire progresser la connaissance sur le sujet étudié. Il est destiné à être publié dans des revues à comité de lecture (peer review) et à ce titre répond aux exigences académiques. Les textes proposés peuvent être en français ou en anglais. En début de texte doivent figurer : les auteurs et leur affiliation institutionnelle, un résumé et des mots clefs.

Le manuscrit sera adressé à : liepp@sciencespo.fr

Les opinions exprimées dans les articles ou reproduites dans les analyses n'engagent que leurs auteurs.

Directeur de publication:
Bruno Palier

Comité de rédaction :
Andreana Khristova, Juliette Seban

Sciences Po - LIEPP
27 rue Saint Guillaume
75007 Paris - France
+33(0)1.45.49.83.61
liepp@sciencespo.fr